Лекция 9

3. ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

ПОСТАНОВКА ЗАДАЧИ

Пусть дано нелинейное уравнение

$$f(x) = 0, \tag{3.1}$$

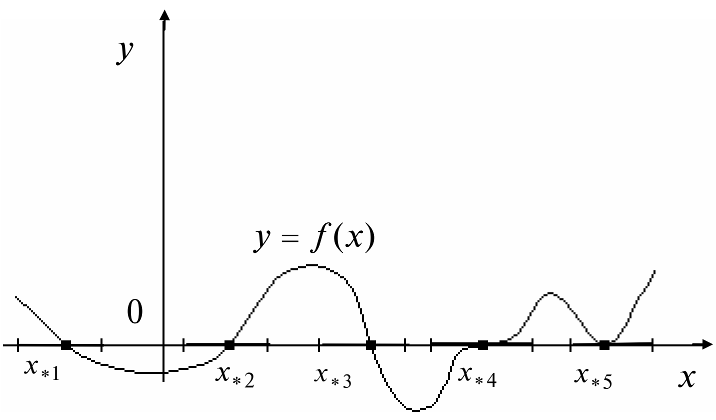
где $f(x)$ – функция, определенная и непрерывная на некотором промежутке. В некоторых случаях на функцию $f(x)$ могут быть наложены дополнительные ограничения, например, непрерывность первой и второй производных, что специально оговаривается. Функция $f(x)$ может быть задана в виде алгебраического многочлена или трансцендентной функции (тогда ей соответствует алгебраическое или трансцендентное уравнение).

Требуется найти *корни* уравнения, т.е. числа $x_{*1}, x_{*2}, \ldots$, которые путем подстановки превращают уравнение в верное числовое равенство. Числа $x_{*1}, x_{*2}, \ldots$ называются также *нулями* функции $f(x)$.

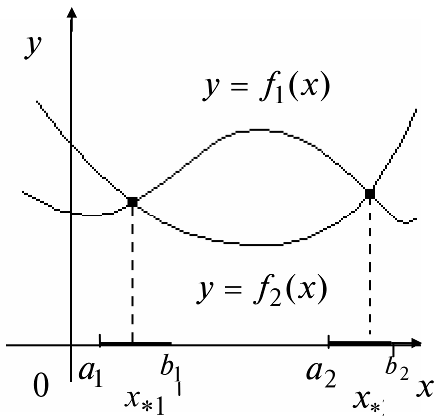
На практике часто бывает выгодно уравнение (3.1) заменить равносильным ему уравнением (уравнения равносильны, если имеют одинаковые корни):

$$f_1(x) - f_2(x) = 0, \tag{3.2}$$

где функции $f_1(x)$, $f_2(x)$ – более простые, чем функция $f(x)$. Тогда при задании уравнения в виде (3.1) нулями функции $f(x)$ являются точки пересечения $f(x)$ с осью Ox (рис.1,*а*), а при задании в виде (3.2) – абсциссы точек пересечения функций $f_1(x)$ и $f_2(x)$ (рис. 1,*б*).



а



б

Рис. 1

Число x_* есть *корень уравнения* (3.1) *кратности* k, если при $x = x_*$ вместе с функцией $f(x)$ обращаются в нуль ее производные до $(k-1)$-го порядка включительно, т.е. $f(x_*) = f'(x_*) = \ldots = f^{(k-1)}(x_*) = 0$, а $f^{(k)}(x_*) \neq 0$. Корень кратности $k = 1$ называется *простым*. На рис. 1,*а* простыми корнями являются x_{*1}, x_{*2}, x_{*3}, а корни x_{*4}, x_{*5} – кратные.

З а м е ч а н и я.

1. Если $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = P_n(x)$ – алгебраический многочлен, то уравнение (3.1) называется также *алгебраическим* n-й степени:

$$P_n(x) \equiv a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0, \tag{3.3}$$

где $a_n, \ldots a_0$ – действительные числа, коэффициенты уравнения.

2. Алгебраическое уравнение n-й степени имеет ровно n корней, действительных или комплексных, при условии, что каждый корень считается столько раз, какова его кратность.

3. Если функция $f(x)$, определяющая уравнение $f(x) = 0$, на концах отрезка $[a_i, b_i]$ принимает значения разных знаков, т.е. $f(a_i) \cdot f(b_i) < 0$, то на этом отрезке содержится по крайней мере один корень уравнения. Если же $f(x)$ непрерывна и дифференцируема и ее первая производная сохраняет знак внутри отрезка $[a_i, b_i]$ $\left(\underset{[a_i, b_i]}{\operatorname{sign}} f'(x) = \mathrm{const}\right)$, то на $[a_i, b_i]$ находится только один корень x_{*i} уравнения.

Этапы решения нелинейных уравнений

Первый этап. Находятся отрезки $[a_i, b_i]$, внутри каждого из которых содержится один простой или кратный корень $(x_{*i} \in [a_i, b_i])$ (см. рис.1). Этот этап называется процедурой *отделения корней*. По сути, на нем осуществляется грубое нахождение корней x_{*i}.

Второй этап. Грубое значение каждого корня x_{*i} уточняется до заданной точности одним из численных методов, в которых реализуются последовательные приближения.

Отделение корней

Для отделения действительных корней полезно определять заранее число корней, а также верхнюю и нижнюю границы их расположения.

В вычислительной практике обычно используются следующие *способы отделения корней*:

1) средствами машинной графики: функция $f(x)$ представляется на дисплее и приближенно определяются отрезки, которым принадлежат точки x_{*i};

2) средствами математического анализа с помощью исследования функций и построения графиков (см. рис. 1,*а*);

3) формированием простых функций $f_1(x)$ и $f_2(x)$ таких, что получается равносильное уравнение в виде (3.2), и дальнейшим построением графиков этих функций (см. рис. 1,*б*).

А. МЕТОД ПРОСТЫХ ИТЕРАЦИЙ

Пусть известно, что корень x_* уравнения $f(x)=0$ лежит на отрезке $G=\{a \le x \le b\}$.

Методика решения задачи

Шаг 1. Уравнение $f(x)=0$ равносильным преобразованием привести к виду $x=\varphi(x)$. Это преобразование может быть осуществлено различными путями, но для сходимости нужно обеспечить выполнение условия $\left|\varphi'(x)\right| \le \chi < 1$ (χ – некоторая константа). При этом задача сводится к нахождению абсциссы точки пересечения прямой $y=x$ и кривой $y=\varphi(x)$ (рис. 2).

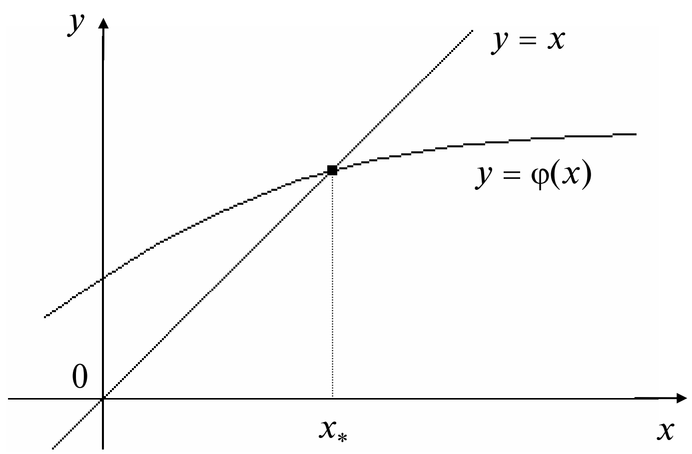


Рис. 2

Шаг 2. Задать начальное приближение $x^{(0)} \in [a,b]$ и малое положительное число ε. Положить $k=0$.

Шаг 3. Вычислить следующее приближение:

$$x^{(k+1)} = \varphi(x^{(k)}) .$$

Шаг 4. Если $\left|x^{(k+1)} - x^{(k)}\right| \le \varepsilon$, итерации завершаются и $x_* \cong x^{(k+1)}$. Если $\left|x^{(k+1)} - x^{(k)}\right| > \varepsilon$, положить $k = k+1$ и перейти к п.3.

Теорема (о достаточном условии сходимости метода простых итераций).

Пусть выполнены условия:

1. Функция $\varphi(x)$ *имеет производные для всех* $x \in G$.

2. Существует число χ $(0 \le \chi < 1,\ \chi = \mathrm{const})$, *такое, что* $\left|\varphi'(x)\right| \le \chi$ *для всех* $x \in G$.

Тогда последовательность $x^{(0)}, x^{(1)}, \dots, x^{(k+1)}, \dots$, *определяемая на основе итерационного процесса, сходится к решению* x_*, *т.е.* $x^{(k)} \to x_*$ *при* $k \to \infty$.

Геометрическая интерпретация процесса сходимости и расходимости в зависимости от выполнения или невыполнения достаточного условия сходимости представлена на рис. 3. Из рис. 3 видно, что при $0 < \varphi'(x) < 1$ и при $-1 < \varphi'(x) < 0$ (см. рис. 3, *а, б*) итерационные последовательности $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$ сходятся к x_*. Причем в первом случае реализуется односторонняя (монотонная) сходимость, а во втором – двусторонняя (немонотонная). При $\left|\varphi'(x)\right| > 1$ (см. рис. 3, *в, г*) процесс расходится, несмотря на то, что точка $x^{(0)}$ очень близка к x_*.

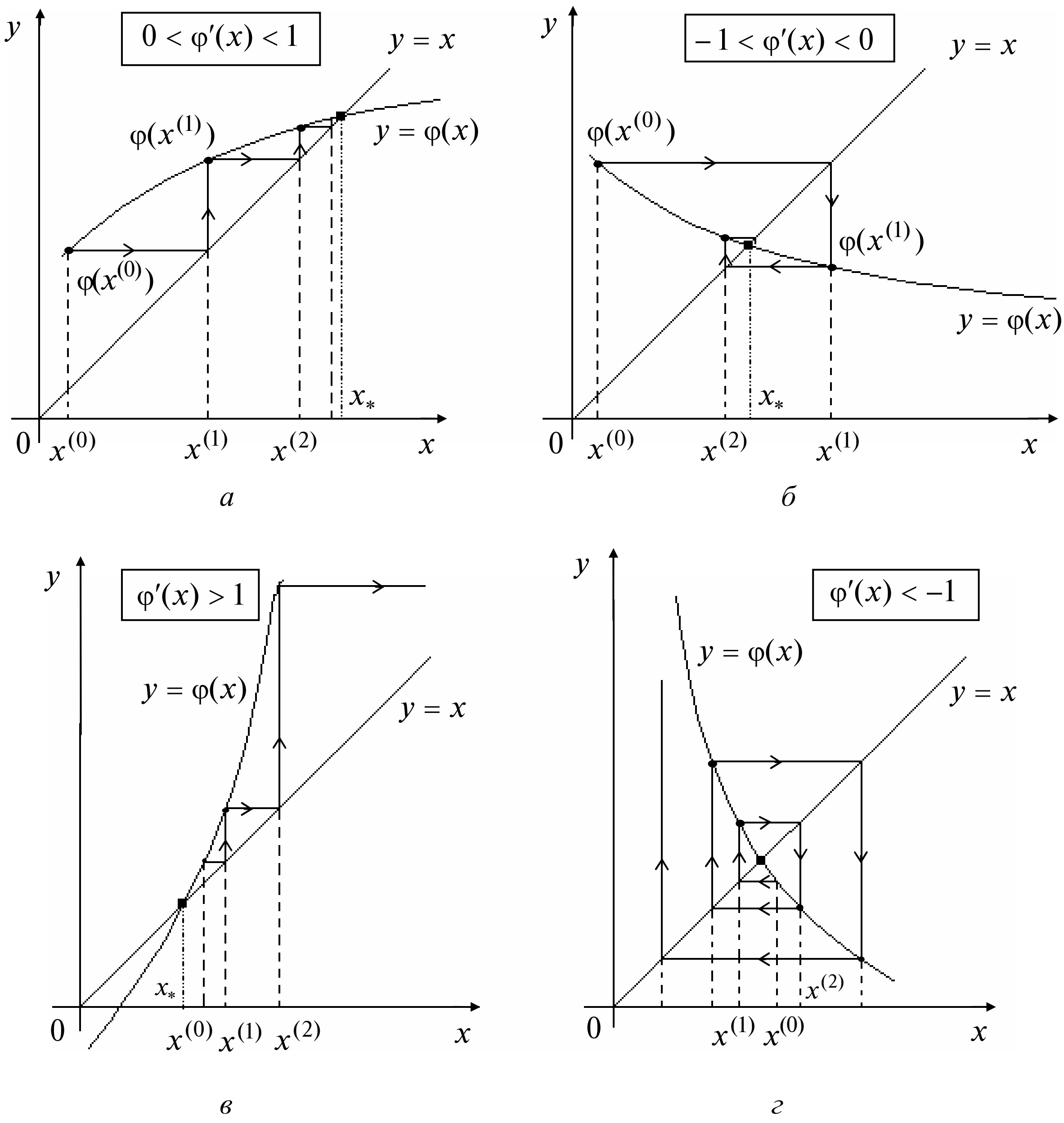


Рис. 3

Способы преобразования уравнения

Преобразование уравнения $f(x)=0$ к равносильному виду $x=\varphi(x)$ может быть выполнено неоднозначно.

1. Можно заменить уравнение $f(x)=0$ на равносильное $x=x+cf(x)$, где $c=\mathrm{const}\neq 0$. Тогда, принимая правую часть этого уравнения за $\varphi(x)$ и раскрывая $|\varphi'(x)|=|1+cf'(x)|<1$, получаем условие

$$-2<cf'(x)<0.$$

При этом надо стремиться получить такую постоянную c, которая бы больше отличалась от нуля, и тогда будет реализовываться более быстрая сходимость.

2. Уравнение $f(x)=0$ заменяется равносильным:

$$x=x\mp\frac{f(x)}{\max|f'(x)|}\equiv\varphi(x) \text{ при } x\in G,$$

где знак в правой части выбирается из условия $|\varphi'(x)|<1$.

3. Можно выразить x из уравнения $f(x)=0$ так, чтобы для полученного уравнения $x=\varphi(x)$ выполнялось условие сходимости $|\varphi'(x)|<1$ в окрестности искомого корня.

Б. МЕТОД НЬЮТОНА

Метод Ньютона (*метод касательных*) является одним из наиболее популярных численных методов. Он реализуется по формуле:

$$x^{(k+1)}=x^{(k)}-\frac{f(x^{(k)})}{f'(x^{(k)})}, \quad k=0,1,2,...$$

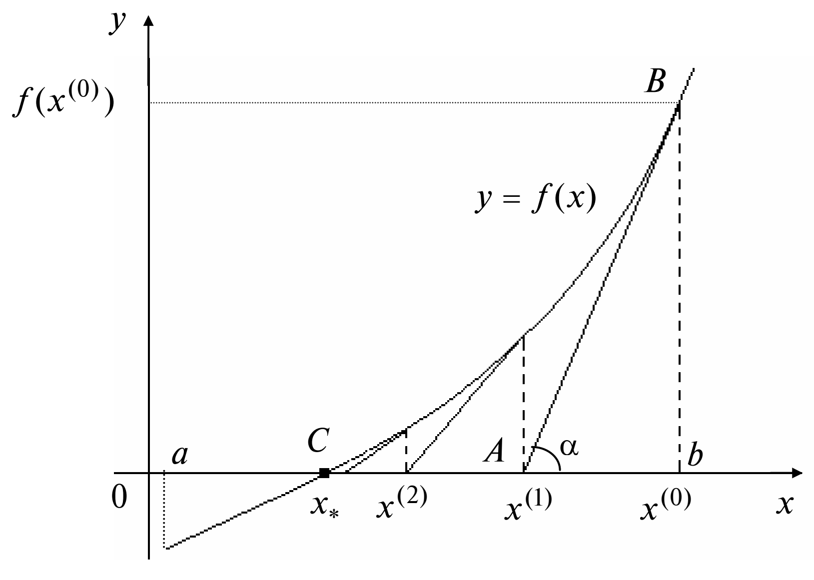


Рис. 4

В точке $x^{(0)}$ строится касательная к графику функции. Следующей точкой $x^{(1)}$ является точка пересечения касательной с осью абсцисс. Далее процесс продолжается аналогично.

Теорема (о достаточных условиях сходимости метода Ньютона).

Пусть выполняются следующие условия:

1. Функция $f(x)$ определена и дважды дифференцируема на $[a,b]$.

2. Отрезку $[a,b]$ принадлежит только один простой корень x_*, так что $f(a)\cdot f(b)<0$.

3. Производные $f'(x), f''(x)$ на $[a,b]$ сохраняют знак, и $f'(x)\neq 0$.

4. Начальное приближение $x^{(0)}$ удовлетворяет неравенству $f(x^{(0)})\cdot f''(x^{(0)})>0$ (знаки функций $f(x)$ и $f''(x)$ в точке $x^{(0)}$ совпадают).

Тогда с помощью метода Ньютона можно вычислить корень уравнения $f(x)=0$ с любой точностью.

В. МОДИФИКАЦИИ МЕТОДА НЬЮТОНА

В1. **Упрощенный метод Ньютона.** Вместо формулы метода Ньютона использует-ся

$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f'\left(x^{(0)}\right)}, \quad k = 0,1,...$$

Отличие от метода Ньютона заключается в том, что производная функции $f(x)$ подсчитывается только в точке начального приближения, а на последующих итерациях не уточняется. Процесс последовательных приближений отражен на рис. 5. Первая итерация совпадает с первой итерацией метода Ньютона. На последующих итерациях соответствующие отрезки параллельны касательной, проведенной в начальной точке.

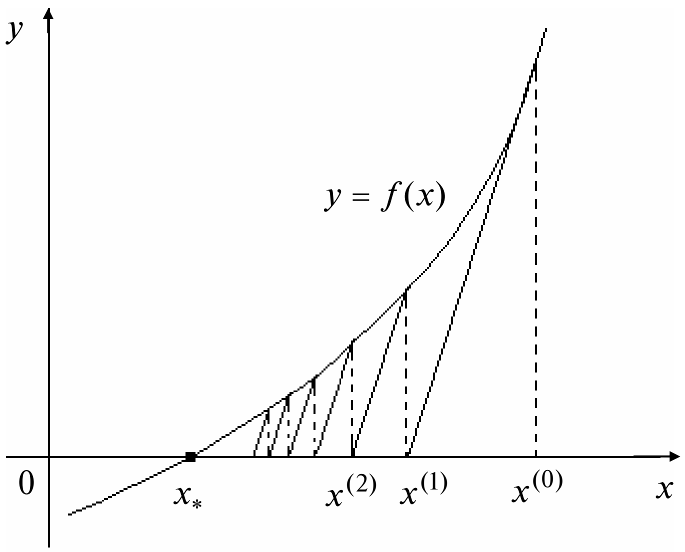


Рис. 5

В2. Метод Ньютона–Бройдена. Этот метод позволяет увеличить скорость сходимости последовательных приближений благодаря использованию формулы

$$x^{(k+1)} = x^{(k)} - c_k \frac{f\left(x^{(k)}\right)}{f'\left(x^{(k)}\right)}, \quad k = 0,1,\ldots,$$

где c_k – число, которое выбирается на каждой итерации так, чтобы уменьшить значение $\left|f\left(x^{(k+1)}\right)\right|$ по сравнению с $\left|f\left(x^{(k)}\right)\right|$. При $c_k = 1$ метод Ньютона–Бройдена совпадает с методом Ньютона.

Как правило, при плохой сходимости или ее отсутствии полагают $0 < c_k < 1$, а при хорошей сходимости для $c_k = 1$ полагают $c_k > 1$ (это ускоряет сходимость).

В3. Метод секущих. В этом методе производная функции $f(x)$ подсчитывается с помощью конечно-разностных соотношений:

– в точке $x^{(0)}$ используется формула $f'\left(x^{(0)}\right) \approx \dfrac{f\left(x^{(0)}\right) - f\left(x^{(0)} - \delta\right)}{\delta}$,

где δ – малая положительная величина;

– в точках $x^{(k)}$, $k = 1,2,\ldots$, используется формула $f'\left(x^{(k)}\right) \approx \dfrac{f\left(x^{(k)}\right) - f\left(x^{(k-1)}\right)}{x^{(k)} - x^{k-1}}$.

Вычисленное значение $f'(x^{(k)})$ определяет тангенс угла наклона секущей (рис. 6).

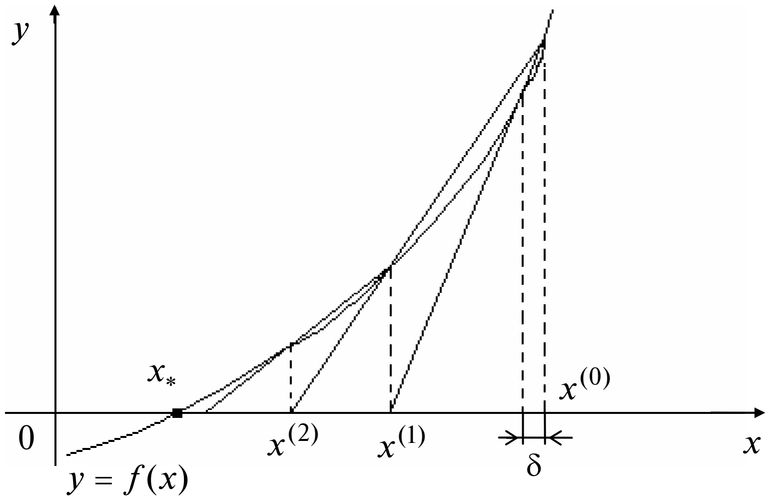


Рис. 6

Используется формула

$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f\left(x^{(k)}\right) - f\left(x^{(k-1)}\right)} \cdot \left(x^{(k)} - x^{(k-1)}\right), \quad k = 1,2,\ldots$$

Г. МЕТОД ПОЛОВИННОГО ДЕЛЕНИЯ

Пусть дано уравнение $f(x) = 0$ и отделен простой корень x_*, т.е. найден такой отрезок $[a_0, b_0]$, что $x_* \in [a_0, b_0]$, и на концах отрезка функция имеет значения, противоположные по знаку ($f(a_0) \cdot f(b_0) < 0$). Отрезок $[a_0, b_0]$ называется *начальным интервалом неопределенности*, потому что известно, что корень ему принадлежит, но его местоположение с требуемой точностью не определено.

Процедура уточнения положения корня заключается в построении последовательности вложенных друг в друга отрезков, каждый из которых содержит корень уравнения. Для этого находится середина текущего интервала неопределенности $c_k = \dfrac{a_k + b_k}{2}$, $k = 0,1,\ldots$, и в качестве следующего интервала неопределенности из двух возможных выбирается тот, на концах которого функция $f(x)$ имеет разные знаки (рис. 7).

Процесс завершается, когда длина текущего интервала неопределенности становится меньше заданной величины ε, задающей точность нахождения корня. В качестве приближенного значения корня берется середина последнего интервала неопределенности.

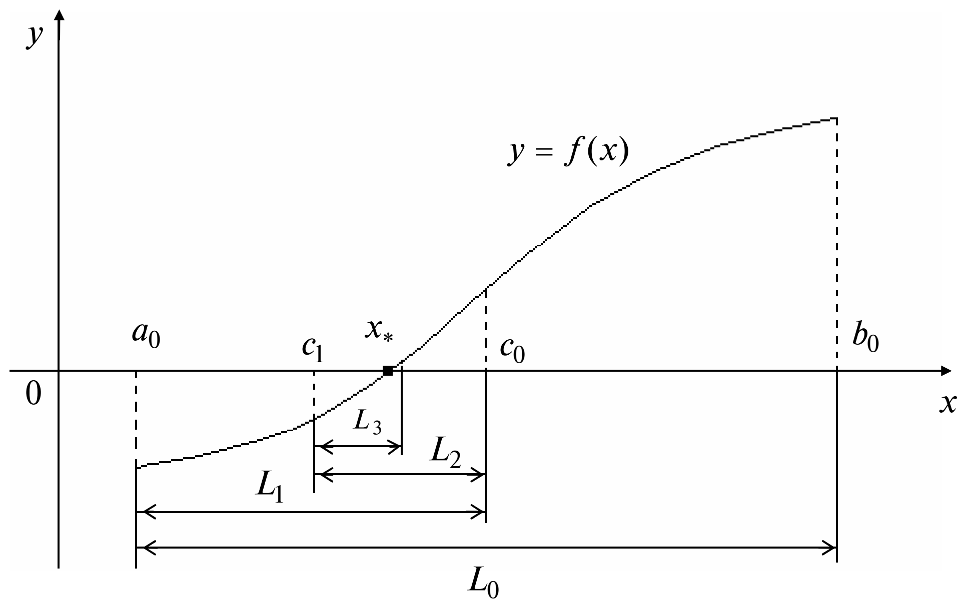


Рис. 7

Д. МЕТОД ХОРД

Этот метод при тех же предположениях обеспечивает более быстрое нахождение корня, чем метод половинного деления. Для этого отрезок $[a, b]$ делится не пополам, а в отношении $|f(a)| : |f(b)|$.

Геометрически метод хорд эквивалентен замене кривой $y = f(x)$ хордой, проходящей через точки $(a, f(a))$ и $(b, f(b))$ (рис. 8).

Уравнение хорды AB имеет вид $\dfrac{x-a}{b-a} = \dfrac{y - f(a)}{f(b) - f(a)}$. Полагая $x = x^{(1)}$ и $y = 0$, получаем $x^{(1)} = a - \dfrac{f(a)}{f(b) - f(a)}\,(b-a)$.

Предположим, что вторая производная $f''(x)$ сохраняет постоянный знак, и рассмотрим два случая: $f(a) > 0$, $f''(x) > 0$ (рис. 9,*а*) и $f(a) < 0$, $f''(x) > 0$ (рис. 9,*б*). Случай $f''(x) < 0$ сводится к рассматриваемому, если уравнение записать в форме: $-f(x) = 0$.

Первому случаю (см. рис. 9,*а*) соответствует формула

$$x^{(0)} = b,$$
$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f\left(x^{(k)}\right) - f(a)} \left(x^{(k)} - a\right), \quad k = 0,1,..., \tag{I}$$

а второму случаю (см. рис. 9,*б*) :

$$x^{(0)} = a,$$
$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f(b) - f\left(x^{(k)}\right)} \cdot \left(b - x^{(k)}\right), \quad k = 0,1,... \tag{II}$$

В первом случае остается неподвижным конец a, а во втором случае - конец b.

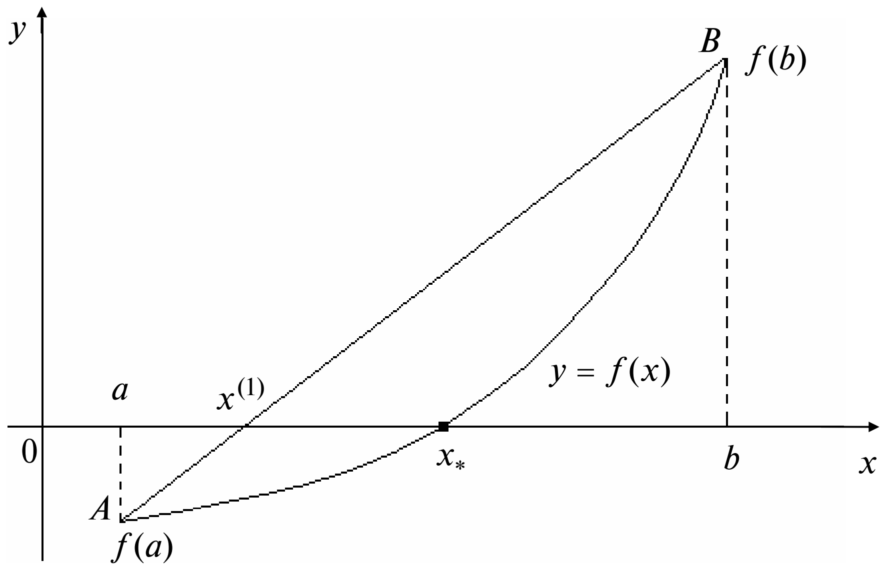


Рис. 8

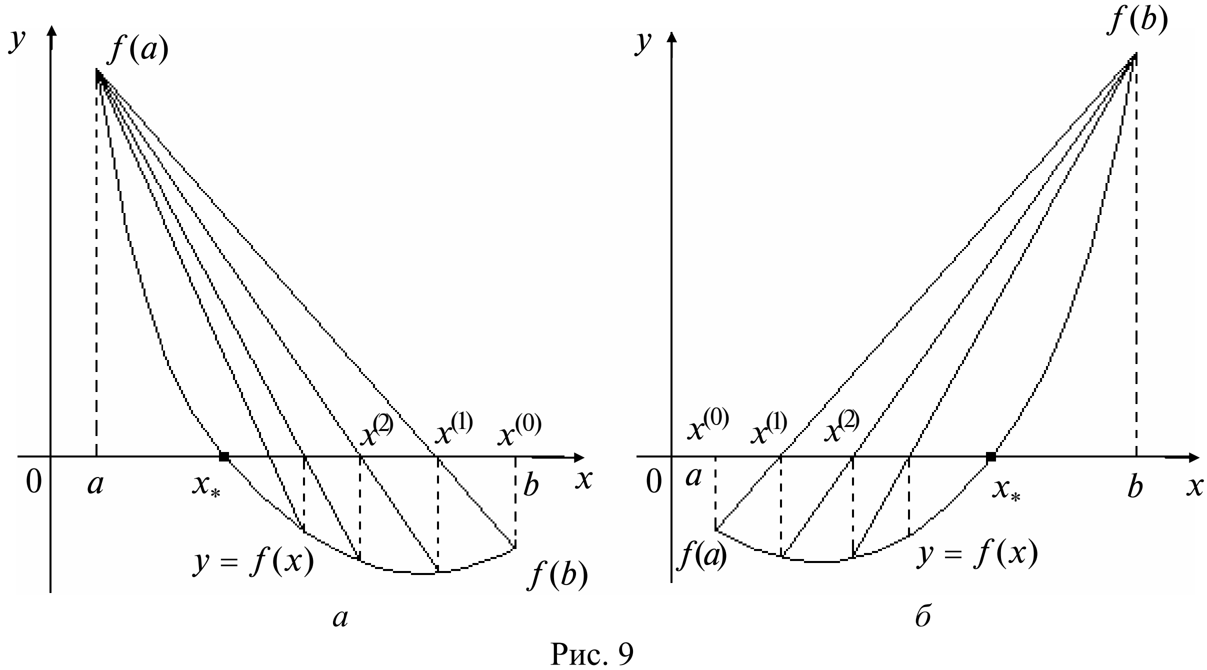


Рис. 9

З а м е ч а н и е. Для выявления неподвижного конца используется условие $f''(x) \cdot f(t) > 0$, где $t = a$ или $t = b$. Если неподвижен конец *a*, применяется формула (I), а если конец b , – формула (II).